Knowledge-Concealing Evidencing of Knowledge about a Quantum State

Emily Adlam
Centre for Quantum Information and Foundations, DAMTP, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge, CB3 0WA, U.K.

Adrian Kent
Centre for Quantum Information and Foundations, DAMTP, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge, CB3 0WA, U.K. and Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, ON N2L 2Y5, Canada.
(Dated: November 29, 2017)

Bob has a black box that emits a single pure state qudit which is, from his perspective, uniformly distributed. Alice wishes to give Bob evidence that she has knowledge about the emitted state while giving him little or no information about it. We show that zero-knowledge evidencing of such knowledge is impossible in quantum relativistic protocols, extending a previous result of Horodecki et al. We also show that no such protocol can be both sound and complete. We present a new quantum relativistic protocol which we conjecture to be close to optimal in security against Alice and which reveals little knowledge to Bob, for large dimension d. We analyse its security against general attacks by Bob and restricted attacks by Alice.

INTRODUCTION

Zero-knowledge proving is a cryptographic primitive in which one agent proves a fact to another agent without giving away any information other than that the fact is true. It has a wide range of practical applications, particularly in electronic voting schemes [2] and digital signature schemes [3], and is also used for a variety of theoretical purposes, such as showing that a language is easy to prove [4]. Zero-knowledge proving of *knowledge*, where Alice is required to prove only that she knows some fact, without giving Bob any information about the fact itself, is a particularly useful version of this task which plays a key role in a number of identification protocols [5].

Horodecki et al. [6] explored the possibility of what they called a "zero knowledge convincing protocol on quantum bit". In their model, a verifier (henceforth called Bob) knows he has a single copy of a pure qubit, but has no other information about the state. A prover (henceforth called Alice) wishes to make a prediction that Bob can verify and that will hold with certainty only if she knows what the state is, without giving Bob any additional information about its identity. They showed that no non-relativistic protocol involving classical information exchanges and quantum Alice-to-Bob communications can implement this task securely [6]. They also discussed some protocols that implement very weak versions of the task, either giving Bob a great deal of information about the qubit, or giving him only weak evidence of Alice's knowledge, or both.

Horodecki et al.'s pioneering discussion was informal on some points. It did not fully distinguish cases in which Alice has classical knowledge about Bob's quantum state (e.g. a classical data string describing it) from cases in which she has quantum knowledge (e.g. a box able to make only some fixed number of copies). Nor did it underline that Alice cannot *prove* that she knows a precise classical description of a single quantum state even if she is not concerned about giving Bob information. This is because the classical information about the state that can be extracted by measurement is bounded, and Alice always has a boundedly nonzero chance of guessing this information even if she knows nothing about the state. (See Theorem 2 below.) Alice may also have a high chance of guessing the information even if she has only partial information about the quantum state – for example, she can predict the outcome of a complete projective measurement on a qubit with probability $\frac{1}{2}$. Similarly, even if she only knows a dimension 2 subspace in which a qudit lies, she can still specify a complete projective measurement whose outcome she can predict with probability $\frac{1}{2}$. Moreover, if the state is η and Alice believes it is η', where the fidelity $F(\eta, \eta')$ is close to 1, then she is almost as likely to pass any protocol testing her knowledge of η as she would be if she knew η. Since Bob only has a single copy of the state, Alice cannot provide more evidence by repeating a protocol that tests her knowledge. By contrast, in classical contexts Alice's chances of success can typically be made arbitrarily small by iteration, so that one can reasonably (modulo epsilonics) speak of classical zero knowledge *proofs*.

Other interesting questions left open include: What can relativistic protocols achieve? How much evidence can Alice provide? What bounds exist on the tradeoffs between the evidence Alice provides and the amount of knowledge she gives away? Are there protocols strong enough for practical cryptographic purposes? How do the answers depend on the dimension d of the state space?

We explore these questions below, beginning with some formal definitions. We prove a stronger no-go theorem showing that no protocol that provides non-trivial evidence of Alice's knowledge about a pure quantum state of finite dimension can prevent Bob from acquiring some additional knowledge about the state, even in the setting of relativistic quantum cryptography. We also prove a bound on the strength of evidence Alice can provide. Since proofs of knowledge of a finite-dimensional quantum state are not possible, and zero-

knowledge protocols that give some evidence of knowledge are also not possible, we then consider the weaker but feasible task of *knowledge-concealing evidencing of knowledge about a quantum state* (KCEKQS).

A KCEKQS protocol requires Alice to give Bob evidence that she has some form of knowledge about a quantum state whilst giving him incomplete information about the state. Ideally, a successfully completed protocol should give Bob as much evidence as possible, without assuming Alice's honesty. Ideally, too, the protocol should ensure as small a bound as possible on the information obtainable by Bob, whether or not it is successfully completed or he honestly follows it. We discuss some simple protocols, generalising protocols previously considered by Horodecki et al. [6], and show that they are relatively weak in knowledge-concealment, or in evidencing knowledge, or both. We then propose a new relativistic quantum protocol. We show it is secure against restricted attacks by Alice and general attacks by Bob, for large d, in a sense we make precise below. We conjecture this remains true for general attacks by Alice.

DEFINITIONS

We assume for now that Alice has no option to abort the protocol; allowing an abort option does not significantly change our main results [1].

A non-relativistic *knowledge-concealing evidencing of knowledge about a quantum state (KCEKQS)* protocol involves two mistrustful parties, Alice and Bob, occupying disjoint secure laboratories. We assume that each party has trusted error-free devices in their own laboratory; both parties trust error-free classical and quantum communication channels between the laboratories; we consider errors and losses later [1]. Bob begins in possession of a quantum system Q_B known to be prepared in some pure state $\eta = |\eta\rangle\langle\eta|$ drawn uniformly at random from Q_B. The protocol requires Alice and Bob to act on alternate rounds and terminates after a fixed finite number of rounds. Each round may require a party to carry out unitary operations and/or measurements on a quantum system in their possession and/or to send classical and/or quantum information to the other party. The protocol may specify that these actions are probabilistically determined, according to given probability distributions. The final round of the protocol requires Bob to generate one of two possible outcomes, 0 and 1, from the classical and quantum information in his possession. These correspond to Bob's rejecting or accepting that Alice has provided evidence of knowledge about η. We write $p(0)$ and $p(1)$ for the outcome probabilities.

In a relativistic KCEKQS protocol, each party may have several trusted agents occupying separate secure laboratories, with secure communications between them, lying within preagreed regions. One agent of Bob's initially possesses Q_B. The protocol requires Alice's and Bob's agents to carry out unitary operations and/or measurements on quantum systems in their possession and to send classical and/or quantum communications to given other agents of the same party and/or the other party, within their agreed location regions and within agreed time intervals. The protocol may specify these actions are probabilistically determined, according to given probability distributions. The protocol terminates after a fixed finite number of such actions. The final prescribed action is for one of Bob's agents to generate one of two possible outcomes, 0 and 1, from the classical and quantum information in his possession, as above.

We will characterise the efficiency of an ideal KCEKQS protocol by three parameters ϵ_{C}, ϵ_{K} and ϵ_{S}; we discuss other relevant features of KCEKQS protocols in the supplementary material [1]. When evaluating these parameters for specific protocols, we will mostly consider the ideal case of error-free devices and channels. In realistic implementations, channel noise, device errors and losses may alter the parameter values.[14] However, our no-go theorems below still hold in reasonable models of noise, errors and losses, so long as these are uncorrelated with η and with any knowledge Alice may have of or about η.

We define our parameters by the following criteria, in each case averaging over η:

- **Completeness:** If Alice has a precise classical description of η and both parties perform the protocol correctly, then $p(1) = 1 - \epsilon_{\mathrm{C}}$.

- **Soundness:** If Alice has no classical or quantum information about the state η, then ϵ_{S} is the supremum of $p(1)$ over all possible (honest or dishonest) strategies for Alice, assuming that Bob performs the protocol correctly.

- **Knowledge-concealing:** Suppose that Alice performs the protocol correctly and at the start of the protocol Bob knows nothing about the state η. Then ϵ_{K} is the supremum of the expected squared fidelity $F^2(\eta, \phi) = |\langle\eta|\phi\rangle|^2$ over (honest or dishonest) strategies that give Bob the value of a pure state ϕ as a guess for η.

Let ϵ_{M} be the supremum of the same expected squared fidelity obtainable by Bob if he does not take part in the protocol and carries out quantum operations and measurements on Q_B. We call $\epsilon_{\mathrm{K}} - \epsilon_{\mathrm{M}}$ the *knowledge gain* available from the protocol to a dishonest Bob. We say the protocol is *zero-knowledge* if $\epsilon_{\mathrm{K}} = \epsilon_{\mathrm{M}}$, We say it is *non-trivial* if $1 - \epsilon_{\mathrm{C}} > \epsilon_{\mathrm{S}}$.

As defined above, general protocols allow both classical and quantum communications in both directions. We will also consider examples with more restricted communications. Extending the discussion of Ref. [6], we consider *classical* protocols, in which Alice and Bob employ only classical communication, *quantum A-to-B* protocols, which additionally allow quantum communications from A to B, and similarly *quantum B-to-A* protocols.

NO-GO THEOREMS

Horodecki et al. [6] showed that no non-relativistic KCEKQS classical or quantum A-to-B protocol for an unknown qubit has $\epsilon_{\rm C} = 0$, $\epsilon_{\rm S} < 1$, and is zero-knowledge. We establish here a considerably more general result, applying to relativistic KCEKQS protocols for qudits with two-way classical and/or quantum communications, and general parameter values. Our statements apply to protocols whose security is based only on quantum theory and special relativity, i.e. within the standard scenario for unconditionally secure relativistic quantum cryptography [8].

Theorem 1. *There exists no non-trivial zero-knowledge KCEKQS protocol. [1]*

The proof of this theorem depends on the fact that in any KCEKQS protocol, Bob may always attempt to cheat by replacing $Q_{\rm B}$ with a system prepared in a state known to him. In heuristic terms, for a non-trivial protocol the probability that Alice nonetheless produces a correct proof is higher if Bob fortuitously chooses a state which is close in Hilbert space to η, and hence Alice's success or failure gives Bob non-zero information about the identity of the unknown state. Since he has retained a copy of the unknown state he can additionally perform a measurement on it, and combining information from these two processes gives him on average strictly more information than is available from the measurement alone. In the supplementary information we formalize this argument and generalize it to cover the case where Alice is allowed to abort the protocol.

We also establish a relationship between completeness and soundness which bounds the degree of evidence Alice can provide:

Theorem 2. *For any qudit KCEKQS protocol,* $\frac{\epsilon_{\rm S}}{1-\epsilon_{\rm C}} \geq \frac{1}{d}$. *[1]*

The proof of this theorem depends on the fact that in any KCEKQS protocol, if Alice does not in fact know the state η, she may always attempt to cheat by choosing a random state ϕ from the Hilbert space of $Q_{\rm B}$ and proceeding with the protocol as if she knows that $Q_{\rm B}$ is in the state ϕ. We show in the supplementary information that the probability of success by this strategy is lower bounded by $\frac{1}{d}(1-\epsilon_{\rm C})$. Hence in any KCEKQS protocol there is inevitably a tradeoff between minimizing the probability that a dishonest Alice manages to cheat and maximizing the probability that an honest Alice manages to produce a successful proof.

In particular, theorem 2 means that for small d, $\epsilon_{\rm S}$ and $\epsilon_{\rm C}$ cannot both be close to 0, regardless of the value of $\epsilon_{\rm K}$. This makes the case of large d particularly interesting to explore.

We observe that the bound of theorem 2 is tight. For example, it is attained by a protocol in which Alice predicts to Bob the outcome of a projective measurement that includes η on the system Q_B: this has $\epsilon_{\rm S} = \frac{1}{d}$ and $\epsilon_{\rm C} = 0$. More generally, it is attained for a protocol in which Alice is required to predict this outcome and also predict the outcome of some independent random event with success probability p: this has $\epsilon_{\rm S} = \frac{p}{d}$ and $\epsilon_{\rm C} = 1-p$. We say a KCEKQS qudit protocol is *CS-optimal* if $\epsilon_{\rm S} = \frac{1}{d}, \epsilon_{\rm C} = 0$.

PROTOCOLS

Classical A-to-B

As noted above, Horodecki et al. [6] argue that non-trivial non-relativistic zero-knowledge classical A-to-B protocols with $\epsilon_{\rm C} = 0$ are impossible for a qubit. In any such protocol, Alice must predict some measurement outcome, and any measurement prediction that holds with certainty for a pure qubit η and is not certain for a random qubit allows Bob to identify η exactly, and so has $\epsilon_{\rm K} = 1$.

One might instead consider protocols with $\epsilon_{\rm C} > 0$, in which Alice chooses a projective measurement which includes a randomly chosen projector P from those with $\langle\eta|P|\eta\rangle = 1-\epsilon_{\rm C}$. One might also consider strengthening such protocols by allowing Alice to use a secure relativistic bit commitment [7–11] to commit her predicted outcome, unveiling this commitment if and only if Bob's reported outcome agrees with her prediction. However, such protocols still have either $\epsilon_{\rm K}$ or $\epsilon_{\rm C}$ large [1].

Quantum A-to-B

Horodecki et al. [6] also consider a protocol where Alice gives Bob a copy of η; as they note, such protocols can achieve $\epsilon_{\rm C} = 0$ and $\epsilon_{\rm K} < 1$. Indeed it is possible to achieve $\epsilon_{\rm K} \ll 1$ for large d. In this sense, for large d, their protocol outperforms the classical A-to-B protocols just discussed. However, one needs to consider the tradeoffs between $\epsilon_{\rm K}, \epsilon_{\rm C}$ and $\epsilon_{\rm S}$. It is also worth highlighting that this protocol requires Alice only to possess quantum information about η rather than classical information. Alice can ensure $p(1) = 1$ even if she only has a black box that will make only a single copy of η and has no other classical or quantum information about η.

We extend the discussion of Ref. [6] by considering a generalisation of their protocol in which Alice gives Bob N copes of η:

1. Alice prepares N systems $\{S_i\}$ in the state η and gives them to Bob.

2. Bob performs a measurement $\{\Pi_S, \mathbb{I} - \Pi_S\}$ where Π_S is the projector onto the symmetric subspace of the joint state space of the $\{S_i\}$ and Q_B.[15]

3. If the result is Π_S, Bob accepts; otherwise he rejects.

If Alice knows η and follows the protocol, Bob will accept, so this protocol achieves $\epsilon_{\rm C} = 0$. We show [1] that $\epsilon_{\rm S} = \frac{1}{N+1} + \frac{N}{d(N+1)}$. So, for $N = 1$, we have $\epsilon_{\rm S} = \frac{1}{2} + \frac{1}{2d}$,

while, for N large, $\epsilon_S \to \frac{1}{d}$, so the protocol tends to CS-optimality in this limit. However, we also show [1] that $\epsilon_K = \frac{N+2}{N+d+1}$, which tends to 1 for large N, while $\epsilon_M = \frac{2}{d+1}$. Thus $\epsilon_K > \frac{1}{d\epsilon_S}$ and $\epsilon_K - \epsilon_M > \frac{d-1}{d+1}\frac{N}{N+2}\frac{1}{d\epsilon_S}$ for all d, N, which are relatively poor tradeoffs. In particular, near CS-optimality ($\epsilon_S \approx \frac{1}{d}$) implies near-zero concealment ($\epsilon_K \approx 1$) and implies significant knowledge gain ($\epsilon_K - \epsilon_M \approx \frac{d-1}{d+1}$).

A quantum B-to-A protocol

We now propose a new KCEKQS protocol that involves quantum B-to-A and two way classical communication. The protocol is relativistic, meaning that 'Alice' and 'Bob' represent two separate networks of collaborating agents distributed in space-time. In our protocol, Alice and Bob each have two agents, A_1, A_2 and B_1, B_2, configured such that $d(A_1, B_1) \approx d(A_2, B_2) \ll d(A_1, B_2) \approx d(A_2, B_1)$[16]

1. Alice and Bob agree in advance on positive integer security parameters N and q.

2. Bob prepares N quantum systems $\{S_i\}$ in states chosen uniformly at random.

3. Bob randomly permutes the systems $\{S_i\}$ and the system Q_B, assigns them all indices from 1 to $N+1$, and then gives all $N+1$ systems, labelled by their indices, to Alice.

4. Alice carries out the projective measurement $\{\eta\,, \mathbb{I}-\eta\}$ on each of the $N+1$ systems that Bob gave her. Write C' for the list of indices for which she obtains outcome η; let $|C'| = q'$. If $q' \leq q$, she forms a list $C = C' \cup D$, where D is a list of $(q-q')$ copies of the dummy index 0.[17]. If $q' > q$, she picks a random size q sublist C of C'.

5. Alice randomly permutes C and then performs q relativistic bit string commitments [1] committing her to each of the indices in the permuted list. Each bit string commitment is set up so that Alice can commit to any index in $\{0, 1, 2, \ldots N+1\}$.

6. Bob tells Alice the index $x \in \{1, \ldots, N+1\}$ that he assigned to Q_B.

7. If $x \in C$, Alice unveils her commitment to that index. Otherwise she announces failure and Bob rejects.[18]

8. If Alice's unveiled commitment is indeed x, Bob accepts. Otherwise he rejects.

In the supplementary information we give an analysis of the security of this protocol. We show that for any possible attack by Bob, $\epsilon_K \leq \frac{4}{d+1}$, while $\epsilon_M = \frac{2}{d+1}$, so $\epsilon_K \to 0$ and $\epsilon_K - \epsilon_M \to 0$ for large d [1], meaning that the protocol is asmptotically secure against Bob. To analyse security against Alice we treat the bit commitment subprotocol as a 'black box' which we assume to be secure under composition[19], meaning that Alice is restricted to strategies in which she commits to q classical values chosen from $\{0, \ldots, N+1\}$ and unveils one of these committed values. We show that under this security assumption, $\epsilon_S \approx \frac{1}{d}$, and if the parameter q is chosen to be $\lceil \frac{N}{d} \rceil$, then in the limit of large N we have $\epsilon_C \to 0$ for $q = \lfloor \frac{N}{d} \rfloor$; hence the protocol asymptotically tends to CS-optimality. A full security analysis requires a complete analysis of general quantum operations Alice could carry out to produce unveiling data; we leave this for future work.

CONCLUSION

We have proven two no-go theorems demonstrating that even in the relativistic setting there is no perfect KCEKQS protocol for quantum states of finite dimension and bounding the evidence Alice can supply. We have also described a new protocol involving quantum Bob-to-Alice communications and relativistic signalling constraints, which appears to achieve a significant improvement on existing protocols for large d. Although it is not zero knowledge for finite d, it reveals little extra information to Bob for large d. We conjecture that it is asymptotically CS-optimal, i.e. offers essentially optimal security against Alice. We anticipate that this protocol may be a valuable quantum cryptographic primitive in contexts where marginal revelations of information to Bob are acceptable.

Acknowledgments This work was partially supported by an FQXi grant and by Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Research and Innovation.

[1] See Supplementary Material for details.

[2] Lipmaa, H; "Secure Electronic Voting Protocols" in *The Handbook of Information Security* vol. 2, H. Bidgoli (ed.), John Wiley and Sons, 2006.

[3] K. Nguyen, F. Bao, Y. Mu, and V. Varadharajan, "Zero-knowledge proofs of possession of digital signatures and its applications," in *Information and Communication Security* (V. Varadharajan and Y. Mu, eds.), vol. 1726 of *Lecture Notes in Computer Science*, pp. 103–118, Springer Berlin Heidelberg, 1999.

[4] L. Fortnow, "The complexity of perfect zero-knowledge," in *Proceedings of the Nineteenth Annual ACM Symposium on Theory of Computing*, STOC '87, (New York, NY, USA), pp. 204–209, ACM, 1987.

[5] U. Feige, A. Fiat, and A. Shamir, "Zero-knowledge proofs of identity," *Journal of Cryptology*, vol. 1, no. 2, pp. 77–94, 1988.

[6] P. Horodecki, M. Horodecki, and R. Horodecki, "Zero knowledge convincing protocol on quantum bit is impossible," *eprint arXiv:quant-ph/0010048*, Oct. 2000.

[7] A. Kent. Unconditionally Secure Bit Commitment. *Physical Review Letters*, 83:1447–1450, August 1999.

[8] A. Kent, "Secure Classical Bit Commitment using Fixed Capacity Communication Channels," J. Cryptology **18** (2005) 313-335.

[9] A. Kent, "Unconditionally secure bit commitment with flying qudits," *New Journal of Physics*, vol. 13, no. 11, p. 113015, 2011.

[10] E. Adlam and A. Kent, "Deterministic relativistic quantum bit commitment," *International Journal of Quantum Information*, vol. 13, p. 1550029, June 2015.

[11] E. Adlam and A. Kent, "Device-independent relativistic quantum bit commitment," *Physical Review A*, vol. 92, p. 022315, Aug. 2015.

[12] S. M. Barnett, A. Chefles, and I. Jex, "Comparison of two unknown pure quantum states," *Physics Letters A*, vol. 307, pp. 189–195, Feb. 2003.

[13] M. Sedlák, M. Ziman, V. Bužek, and M. Hillery, "Unambiguous comparison of ensembles of quantum states," *Physical Review A*, vol. 77, p. 042304, Apr. 2008.

[14] Typically, uncorrected noise, errors and losses will increase ϵ_{C} and decrease ϵ_{K}, and uncorrected losses will decrease ϵ_{S}.

[15] A motivation for this choice is that, given'a system in state $\psi^{\otimes n}$ and another system in state $\phi^{\otimes m}$, for some integers m, n, the measurement $\{\Pi_S, \mathbb{I} - \Pi_S\}$ (i) always gives outcome 1 if $\psi = \phi$ (ii) maximises the probability of outcome 0 if $\psi \neq \phi$, among measurements satisfying (i) [12, 13].

[16] Here $d(x, y)$ represents the spatial distance between the spatial location of x and the spatial location of y, in the joint rest frame of x and y, so that 'Alice' and 'Bob' can employ the relativistic bit commitment protocol set out in ref [8]. This relativistic bit commitment protocol is used as a sub-protocol in the course of the KCEKQS protocol:

[17] This dummy index prevents cheating strategies in which Bob uses the number of Alice's commitments, made at the next step, to extract additional information about the state.

[18] In the ideal error-free case, if Alice knows η precisely and both parties honestly perform the protocol, failure is possible if and only if $q' > q$.

[19] We conjecture that our conclusions will still hold even when this assumption is lifted, but a full analysis must wait upon a better theoretical understanding of the composability of bit commitment protocols in general.